Real-Time Stock Forecasting: Leveraging Live Data and Advanced Algorithms for Accurate Predictions

Chandu Vaidya¹
B.Tech CSE
S.B. Jain Institute of Technology
Nagpur, Maharashtra India, Asia, 441501

Gulrukh Nazneen²
B.Tech CSE
S.B. Jain Institute of Technology
Nagpur, Maharashtra India, Asia, 441501

Nidhi Singh³
B.Tech (CSE)
S.B. Jain Institute of Technology
Nagpur, Maharashtra India, Asia, 441501

Kapil Katariya⁴
B.Tech (CSE)
S.B. Jain Institute of Technology
Nagpur, Maharashtra India, Asia, 441501

Aditya Ramtekkar⁵
B.Tech (CSE)
S.B. Jain Institute of Technology
Nagpur, Maharashtra India, Asia, 441501

Diptanshu Nasare⁶
B.Tech (CSE)
S.B. Jain Institute of Technology
Nagpur, Maharashtra India, Asia, 441501

Diksha Lalmore⁷
B.Tech (CSE)
S.B. Jain Institute of Technology
Nagpur, Maharashtra India, Asia, 441501

Abstract:- This groundbreaking research introduces an innovative stock market prediction methodology that integrates financial modeling, machine learning, and real-time data analysis. Anchored in a deep understanding of stock market dynamics, including economic statistics, company performance, and market sentiment, our approach employs natural language processing (NLP) and predictive modeling to analyze live data for accurate stock price predictions. The method involves collecting and preprocessing a dynamic dataset enriched with financial indicators and historical stock prices. Utilizing Long Short-Term Memory (LSTM) algorithms, our model exhibits an impressive 96% accuracy in forecasting stock movements, showcasing adaptability to diverse market scenarios and responsiveness to economic factors and sentiment shifts. The incorporation of live data proves pivotal in providing timely insights for informed decision-making, establishing our model as a valuable tool for navigating the complexities of the modern financial landscape.

Keywords:- Stock Market, Machine Learning, Natural Language Processing, Prediction.

I. INTRODUCTION

Analyzing stock market [1] trends is an interdisciplinary area of research being undertaken by many disciplines such as Finance, Computer Science, Statistics, and Economics. Stock market is one of the major fields that investors are dedicated to, thus stock market price trend prediction is always a hot topic for researchers from both financial and technical domains. In this research, our objective is to build a state-of-art prediction model for price trend prediction, which focuses on short-term price trend prediction [8]. The author’s Jingyi Shen, M. Omair Shafiq proposed a fine-tuned and customized deep learning prediction.

[8] system along with utilization of comprehensive feature engineering and combined it with LSTM to perform prediction. They also applied optimization techniques, such as principal component analysis (PCA) were also applied in short-term stock price prediction.

On the other hand, Prof. Bharathi H. N., Kalyani Joshi - They used polarity detection algorithm (PCA) [2] for initially labelling news and making the train set. For this algorithm, dictionary-based approach was used. During the years, researchers are not only focused on stock price-related analysis but also tried to analyze stock market transactions such as volume burst risks, which expands the stock market analysis research domain broader and indicates this research domain still has high potential.
While the researchers frequently proposed different neural network solution architectures [8], it brought further discussions about the topic if the high cost of training such models is worth the result or not. Now a days, advanced intelligent techniques based on either technical or fundamental analysis are used for predicting stock market. Particularly, for stock market analysis, the data size is huge and also non-linear. To deal with this variety of data efficient model is needed that can identify the hidden patterns and complex relations [5] in this large data set.

This innovative method aims to capitalize on the vast amount of information available in real-time news sources, which can influence market sentiment and subsequently impact stock prices. In this cutting-edge application of machine learning to stock market prediction, the algorithm is trained on a live dataset containing vast previous years data of machine learning models for stock market prediction.

The aim of the research work is to predict future stock prices based on live data, with the objective of building a predictive LSTM model for forecasting of stocks.

- **Objectives:**
  - **Development of a Static Prediction Model:**
    This involves constructing a system that utilizes LSTM (Long Short-Term Memory) algorithms to forecast outcomes based on historical data.
  - **Live Dataset Extraction from Yahoo Finance:**
    The process involves gathering and wrangling real-time data from Yahoo Finance, which is then prepared for analysis within the prediction model.
  - **Interactive Interface Integration:**
    Implementing a user-friendly interface that allows individuals to engage with the prediction model, enabling them to select a particular stock to access forecasts in an interactive manner.

## II. LITERATURE SURVEY

Stock price trend vaticination is an active exploration area, as more accurate prognostications are directly related to future returns in stocks. thus, in recent times, significant sweats have been put in into developing models that can prognosticate for unborn trend of a specific stock or overall request. utmost of the being ways make use of the specialized pointers. Some of the experimenters showed that there's a strong relationship between news composition about a company and its stock prices oscillations. Following is discussion on former exploration on sentiment analysis of textbook data and different bracket ways. This exploration paper makes a significant donation background and literature review. The assessment of private responses has been a longstanding pursuit, with nearly two decades of exploration devoted to the development and refinement of methodologies. Over this period, different way have been employed to address this challenge.

- **Saloni Mullapudi, Sahitya Mullapudi (April 2019) [1]** – Chancing unborn trend for a stock is a pivotal task because stock trends depend on number of factors. The accuracy of stock price predictions by gathering a large amount of time series data and analyzing it in relation to related news articles, using deep learning models. They built prediction models based on time series forecasting models, such as ARIMA, RNN, and Facebook Prophet So, completely studied this relationship and concluded that stock trend can be prognosticated using newspapers and former price history.

- **Kalyani Joshi, Prof. Bharathi H.N., [2]** – As newspapers prisoner sentiment about the current request, they automate this sentiment discovery and grounded on the words in the newspapers, can get an overall news polarity. The accuracy of the prediction model is 80% and in comparison with news random labelling with 50% of accuracy; the model has increased the accuracy by 30%. However, also they state that this news impact is good in the request, so more chances of stock price go grandly, If the news is positive. And if the news is negative, also it may impact the stock price to go down in trend. used opposition discovery algorithm for originally labelling news and making the train set. Random Forest worked very well for all test cases ranging from 88% to 92% accuracy. Accuracy followed by SVM is also considerable around 86%. Naive Bayes algorithm performance is around 83%.

- **Sai Vikram Kolasani, Rida Assaf, [3]** – In our work, we prognosticate the unborn movement of the United States stock request by assaying the sentiment of Twitter posts related to the Stock request. Next, they use the sentimental analysis of one year of data of tweets that contain the “stock market”, “stocktwits”, “AAPL” keywords, with the goal of predicting the corresponding stock prices of Apple Inc. (AAPL) and the US’s Dow Jones Industrial Average (DJIA) index prices. We show that neural networks perform substantially better than traditional models for stocks’ price prediction. They got near about 82% of accuracy using SVM model.

- **Deeksha Chandola, Akshii Mehta [4]** – The paper presents a deep literacy model that helps the investors comprehend the request’s trading geste This model outperforms previously proposed models that have used news of past weeks and months, with improved accuracy of 65.4%. The frame combines word embedding with intermittent neural network for. The result supports the hypothesis that the information in the news headlines has a short temporal effect on the investors prognosticating stock price directional movement. The model takes a combination of fiscal time series and news captions as input. Compared to other applicable work, the use of mongrel input appreciatively influences the affair. Regarding the textbook representation, it's observed that both good news and bad news induce a change in the stock price. It helps in the macroeconomic analysis of the stock. Hence, the news titles from the day ahead are aligned in a unique case and used for directional soothsaying.
Mehar Vijh, Deeksha Chandola, (Noida 201304, India) [5]–

In this work, Artificial neural Network and Random forest techniques have been utilized for predicting the next day closing price for five companies belonging to different sectors of operation. The historical dataset available on company’s website consists of only few features like high, low, open, close, adjacent close value of stock prices, volume of shares traded etc., which are not sufficient enough. Results show that the best values obtained by ANN model gives RMSE (0.42), MAPE (0.77) and MBE (0.013).

Findings

Initially, we utilized the LSTM algorithm, achieving an impressive 96% accuracy in training our stock market prediction model. Subsequently, we developed a user interface featuring two graphs: a time series graph providing accurate predictions until the current date and a forecasting graph offering precise predictions for the next 2 years, with a provided range for the subsequent 3 years. This professional approach enhances user accessibility and transparency in communicating the model’s capabilities and limitations, positioning our research at the forefront of stock market prediction methodologies.

Proposed Approach

The core of our approach lies in the development of a sophisticated machine learning model, incorporating algorithms like Long Short-Term Memory (LSTM). This model is trained on historical data with feature engineering to discern critical variables influencing stock prices. Additionally, sentiment analysis on news and social media data, facilitated by Natural Language Processing (NLP) tools, is integrated to gauge market sentiment—a crucial factor in stock price movements.

First and foremost, gathering real-time data is essential. Financial data, such as stock prices, trade volumes, economic indicators, and news mood, can be retrieved in real-time for this purpose. For this, news organizations, market data providers, and financial platforms’ APIs are frequently used.

Statistical analysis can then be used to find patterns and trends within the data. While correlation analysis can show connections between stock prices and different economic factors, time series analysis aids in understanding past price movements. This examination.

This approach helps lay the groundwork for a fundamental comprehension of market behavior. Stock price prediction is significantly aided by machine learning models. Understanding the sentiment of the market can be gained through sentiment analysis of news and social media. Textual data can be analyzed to determine the general sentiment of market participants with the aid of Natural Language Processing (NLP) tools.

In the initial phase of our research, we commenced with the acquisition of raw data from live sources encompassing stock prices, trading volumes, economic indicators, and sentiment scores derived from news feeds. Following this, a meticulous process of feature extraction was undertaken to distill pertinent information, including stock prices, trading volumes, and sentiment scores, establishing a comprehensive feature set for subsequent analysis. The identification of original indices and potential expansion of features ensued, facilitating a nuanced understanding of the dataset’s structural composition. Employing Recursive Feature Elimination (RFE) enabled the systematic selection of the most impactful features, while particular emphasis was placed on those exhibiting high weights or importance scores. To further refine the feature set, Principal Component Analysis (PCA) was employed, serving as a pivotal step in dimensionality reduction while retaining essential variance within the high-weighted features. This multistep process laid the foundation for subsequent stages.

Dataset:

In this segment, we delve into the intricacies of our data extraction journey from various public sources, unveiling the meticulous crafting of our final dataset. Given the inherent diversity of stock market-related data, our approach commenced with a rigorous comparison of existing works within the realm of financial research.

Once our data arsenal was assembled, we undertook the crucial task of defining a robust data structure for the dataset. This structural framework served as the backbone for organizing and interpreting the wealth of information at our disposal. Presented below is an exhaustive exploration of our dataset, offering a nuanced understanding of its architecture, data tables within each category, and a detailed segment definition. This holistic exposition provides a unique insight into the intricacies of our data curation process and lays the foundation for the subsequent analyses and findings presented in this research endeavor.

Description of Dataset:

In this research endeavor, we employed the yfinance Python library as a pivotal tool for acquiring our financial dataset. Yfinance facilitated the seamless extraction of diverse financial data directly from Yahoo Finance. This dynamic Python library stands out for its simplicity and convenience, offering an accessible means to access an extensive array of financial metrics. The dataset encompasses essential components such as historical stock prices, trading volumes, and other pertinent indicators for a multitude of companies.

By leveraging yfinance, we ensured a robust and up-to-date dataset, providing a foundation for our predictive model. The incorporation of live news feed data further enriched the dataset, introducing real-time insights and enhancing our analysis of the intricate relationship between market sentiment and stock price movements. This meticulously curated dataset, drawn from yfinance.
III. METHODOLOGY

- Following System Design (Figure 1) is Proposed in this Project to Classify News Articles for Generating Stock Trend Signal.

![Flowchart of Propose System](image)

Methodology is mainly divided into 3 phases. In Phase 1 we started our project with identifying problem in stock market where we analyse that stock market prediction is mainly depend on the previous pattern or trend of the particular stock. We gathered requirement for making a Deep Learning Model which can predict the stock. We decide which technology to be used and proposed our work. Result of phase 1 is to Developed such model which can predict the stock by taking live data. In phase 2, We collect the live stock data by implementing a library named yfinance which take data from yahoo finance site then we processed that data to make used of data in our model. We used LSTM Algorithm to create our model. At the end of phase 2, we have developed the model which can predict the stock. In phase 3, We started making our User Interface in streamlit framework of python Then we implement our model in python pycharm to connect them. As a result of phase 3, we get our desired output which is prediction of stock.

IV. IMPLEMENTATION

- Model Building:

In the initial phase of our study, we constructed a robust model employing the LSTM (Long Short-Term Memory) algorithm using a static dataset. The dataset was meticulously divided into training and testing sets, with a ratio of 60-40 to ensure effective model evaluation. Prior to training, we diligently preprocessed the raw data to render it suitable for the construction and training of our LSTM model.

Subsequently, the model underwent a rigorous training process, during which it learned patterns and relationships inherent in the static stock dataset. This training phase resulted in the attainment of desired outputs, which were further validated on the testing dataset. The outcomes from the testing phase demonstrated that our model effectively met the specified requirements, affirming the successful accomplishment of our primary objective: the construction of a well-performing model using a static stock dataset.

As we move forward, the subsequent stages of our study will involve the introduction of dynamic elements, such as live news feed data, to enhance the predictive capabilities of our model in the ever-changing landscape of the stock market.

- Providing Live Data to Our Model:

Our second objective focuses on equipping our model with the capability to predict stocks in real-time by incorporating live data. To achieve this, we utilized the yfinance Python library, a powerful tool that facilitates the seamless retrieval of financial data from Yahoo Finance. This library offers a user-friendly and efficient means of accessing a diverse range of financial information.

In the process, we integrated our model with the yfinance library, enabling it to dynamically fetch live data directly from the Yahoo Finance website. This integration transforms our model into a real-time application, providing timely and up-to-date information for stock market predictions. The development environment for this real-time
model is PyCharm, a versatile Python IDE, where the integration with yfinance takes place seamlessly.

As we progress, the utilization of live data ensures that our model remains adaptive and responsive to the dynamic nature of financial markets, enhancing its practical utility for real-world stock market predictions.

- **User Interface:**
  In the final stage of our implementation, we aimed to furnish our model with a user-friendly interface, allowing users to effortlessly access stock predictions. To materialize this objective, we leveraged the Streamlit framework—an open-source tool designed for the rapid development and sharing of aesthetically pleasing machine learning and data science web applications.

  The integration of Streamlit proved instrumental in simplifying the implementation of our user interface. The framework's user-friendly features facilitated the creation of an interactive interface, streamlining the user experience. This interface enables users to input the desired stock-predicting company name and the number of years for prediction. The system then provides search results, presenting the predicted stock outcomes. Our ultimate goal with this implementation is to deliver an accessible and interactive interface that empowers users to seamlessly interact with and interpret stock predictions, fostering a user-friendly and informative experience.

  Hence, our model has evolved into a powerful tool for stock prediction, seamlessly blending the prowess of algorithms like Long Short-Term Memory (LSTM), the versatility of libraries such as yfinance for live data integration, and the user-friendly interface afforded by the Streamlit framework. With these components harmoniously integrated, our model stands prepared to make real-time stock predictions while engaging users through an intuitive interface. By marrying cutting-edge technology with user-centric design, our comprehensive solution embodies the intersection of efficiency, accuracy, and accessibility in the realm of stock market prediction.

V. **RESULTS**

Our findings showcase the model's adaptability to various market scenarios, emphasizing its robustness in capturing trends and responding to shifts in economic factors and sentiment. The incorporation of live data, including financial indicators and sentiment analysis, significantly contributes to the model's predictive capabilities, providing investors with timely insights for informed decision-making. These results affirm the potential of our approach in advancing the field of stock market prediction, offering a valuable tool for navigating the complexities of today's dynamic financial landscape.

Here, figure 2 graph encapsulates stock predictions up to the current date, employing data sourced from Yahoo Finance. This time series graph serves as a reliable indicator, offering accurate predictions for diverse companies. By forecasting the performance of various entities leading up to today, it provides a comprehensive visualization of predicted stock trajectories.

The utilization of data extracted from Yahoo Finance enriches the graph with dynamic and up-to-date information, enabling investors to gain valuable insights into predicted trends. This real-time data integration transforms the graph into a practical tool, assisting investors in making informed and strategic decisions. Investors can leverage this resource to comprehend stock market movements and optimize their investment strategies based on the latest forecasts. In essence, the graph serves as a valuable ally for navigating the complexities of the stock market landscape.
The figure 3 showcased above paints a forward-looking picture, projecting trends for multiple companies over the next five years. Driven by data meticulously extracted from Yahoo Finance, this tool emerges as a valuable resource for gaining insights into the anticipated trajectories of various stocks.

Notably, the graph provides precise results for the immediate next 1-2 years, offering a clear snapshot of expected stock movements. Additionally, for the subsequent 3-4 years, the tool furnishes a range within which the stock is predicted to lie. This dynamic feature enhances flexibility and accommodates varying investment strategies.

Users are empowered with the ability to customize their predictions by selecting the number of years for projection, tailoring the tool to meet their specific forecasting needs. In essence, this forecasting graph emerges as a forward-thinking companion, aiding investors in navigating the intricacies of the financial landscape with foresight and strategic acumen.

VI. DISCUSSION

In this section, we discuss and compare the results of our proposed model, other approaches, and the most related works.

- Comparison with related works

- (SVM), Random Forest (RF), RNN and Convolutional Neural Network (CNN).

In the research paper titled "Predicting Stock Movement Using Sentiment Analysis of Twitter Feed with Neural Networks," authored by Sai Vikram Kolasani and Rida Assaf, the authors explore the dynamic realm of stock market prediction through sentiment analysis derived from Twitter feeds. Notably, their methodology leverages the Support Vector Machine (SVM) algorithm to forecast future stock movements, achieving a commendable accuracy rate of 83%. We focus on the utilization of live data for prediction, providing a real-time perspective on stock trends. Moreover, our approach employs the LSTM algorithm, resulting in a remarkable accuracy of 96% during the training phase.

In the research paper titled "Stock Trend Prediction Using News Sentiment Analysis," authored by Kalyani Joshi, Prof. Bharathi H. N., and Prof. Jyothi Rao, the authors explore the realm of predicting current stock trends through sentiment analysis of news articles. Employing the Random Forest algorithm, their model attains a substantial accuracy of 90%. In contrast, our research takes a unique trajectory, concentrating on the prediction of future stock movements using live data.

In the research paper titled "Stock Price Prediction Using News Sentiment Analysis," authored by Saloni Mohan1, Sahitya Mullapudi1, Sudheer Sammeta1, Parag Vijayvergial and David C., the authors adopted a distinct approach by utilizing Recurrent Neural Networks (RNN) for their predictive model. RNNs are well-suited for capturing sequential dependencies in data, making them particularly relevant for time-series prediction tasks such as stock price forecasting. This choice of algorithm aligns with the objective of incorporating the temporal dynamics of news sentiment to enhance the accuracy of stock price predictions. It’s worth noting that our research, in contrast, employs the Long Short-Term Memory (LSTM) algorithm for a similar purpose, showcasing the diverse range of neural network architectures explored within the domain of stock market prediction.

In the research paper titled "Forecasting Directional Movement of Stock Prices using Deep Learning," authored by Deeksha Chandola, Akshit Mehta, and Shikha Singh, the authors conducted a comparative analysis of the Stock Movement Prediction (SMP) model using Recurrent Neural
Networks (RNN) and Long Short-Term Memory (LSTM) architectures. Notably, their study focused on forecasting the directional movement of stock prices. However, it's important to highlight that the data utilized in their analysis did not involve live data. Despite this, their findings demonstrated an accuracy of 65% in predicting the directional movement of stock prices. This study contributes valuable insights into the application of deep learning techniques for stock price forecasting, offering a comparative evaluation of RNN and LSTM models.

As we delve into the comparison with related work, it becomes evident that our research not only builds upon existing knowledge but introduces a sophisticated paradigm that enhances predictive capabilities in the dynamic realm of stock market forecasting. From the previous works, we found the most commonly exploited models for short-term stock market price trend prediction are support vector machine.

VII. CONCLUSION

The conclusion of the study highlights the effectiveness of the proposed approach, which involves the generation of two copies of data, demonstrating a device-independent replication method. This innovative strategy ensures that data remains accessible even if deleted from one device, as it is retained on another. The significance of this approach lies in its potential to instill trust among users, enhance user satisfaction, and mitigate data loss risks. The study underscores the importance of data replication as a potent technique to minimize user waiting times and augment data availability. By efficiently distributing workload across cloud nodes, the proposed approach contributes to improved response times, addressing the crucial aspect of user experience. The conclusion emphasizes the necessity of a replica replacement method, particularly in the context of restricted storage space, to optimize the efficiency of dynamic replica management. Overall, the findings suggest that the outlined approach not only introduces a resilient and reliable data replication system but also acknowledges the practical challenges associated with storage limitations, offering a comprehensive solution for enhancing overall system performance and user satisfaction.

Furthermore, the conclusion underscores the role of the proposed approach in optimizing resource utilization within cloud computing environments. By efficiently distributing workload across cloud nodes, the approach not only enhances response times but also minimizes the risk of resource bottlenecks. This is particularly crucial in dynamic and scalable computing environments where the demand for resources can vary rapidly. The study emphasizes how the strategy aligns with contemporary cloud computing paradigms, making it a relevant and forward-thinking solution for modern data management challenges. The multifaceted advantages of the device-independent replication approach extend beyond immediate concerns of data accessibility and loss prevention. It addresses critical aspects of disaster recovery, resource optimization, and scalability, making it a comprehensive and forward-looking solution in the ever-evolving field of data management within cloud computing environments.

FUTURE SCOPE

The future scope of "Stock Market Prediction Using Live Data" is vast and dynamic. By embracing technological advancements, expanding the scope to global markets, and fostering collaboration, the project has the potential to revolutionize the way investors navigate the complex landscape of financial markets. Stock market prediction can be used in the areas such as: Incorporation of Advanced NLP Techniques, Real-time Market Impact Assessment, Continuous Model Optimization with Reinforcement Learning, Collaboration with Financial Institutions.

The future scope of stock market prediction using live data on live datasets holds immense potential for revolutionizing investment strategies and financial decision-making. Incorporating real-time news data into stock market prediction models can provide a more dynamic and accurate assessment of market trends, enabling investors to make informed decisions. With the advent of advanced machine learning and natural language processing techniques, the analysis of news sentiment, market sentiment, and other relevant factors can be automated, offering investors a comprehensive understanding of the factors influencing stock prices.

Furthermore, the integration of live datasets ensures that the predictive models to changing market conditions in real time. This responsiveness is crucial in the ever-evolving landscape of financial markets where news and events can have an instantaneous impact. The application of deep learning algorithms, coupled with the continuous influx of live data, can enhance the predictive accuracy of these models, making them more reliable for investors seeking to navigate the complexities of the stock market. Additionally, advancements in technology, such as improved data processing speeds and enhanced computing capabilities, will contribute to the scalability of these predictive models. This scalability allows for the analysis of vast amounts of live data, leading to more robust and nuanced predictions. As the field of artificial intelligence continues to evolve, so too will the sophistication of stock market prediction models, potentially unlocking new insights and uncovering previously unnoticed patterns within the data.

REFERENCES


[6]. Jagruti Hota, Sujata Chakravarty, Stock Market Prediction Using Machine Learning Techniques, Faculty of Management Studies, Sri University, Odisha, India.


[12]. Dr YVS Sai Pragathi, M V S Phani Narasimham and Dr B V Ramana Murthy, Analysis and implementation of realtime stock prediction using reinforcement frameworks, CSE Dept, Stanley College of Engineering & Technology for Women, Hyderabad, India

